specification page 8, lines 1-13; and page 23, lines 4-11 as originally filed. No new matter is added. This Amendment is being made to expedite prosecution, and not for reasons related to patentability.

From the Detailed Action:

Specification

The Examiner has objected to the specification because it contains embedded hyperlink and/or other form or browser-executable code. Applicants have adopted the Examiner's suggestion and amended the specification to remove any hyperlinks. Accordingly, Applicants have obviated this objection.

Claim Objections

The Examiner has objected to the claims because Claims 5-10 are Claims 13-18 are substantial duplicates thereof.

Applicants respectfully disagree. Two independent claims exist: Claims 1 and 11. Claim 1 describes the process of "clustering" the datapoints using a self organizing map (SOM), whereas claim 11 refers to "grouping" the datapoints.

The MPEP, section 706.03(k), states:

> Inasmuch as a patent is supposed to be limited to only one invention, or, at most, several closely related indivisible inventions, limiting an application to a single claim, or a single claim to each of the related inventions might appear to be logical as well as convenient. However, court decisions have confirmed applicant's right to ***restate (i.e., by plural claiming) the invention*** in a reasonable number of ways. Indeed, a mere difference in scope between claims has been held to be enough. *(Emphasis added)*

Applicants have chosen to restate the invention in these two ways. Hence, Applicants respectfully request reconsideration and withdrawal of this objection.

Rejection of Claims 1-18 under 35 U.S.C. §103(a)

The Examiner has rejected Claims 1-18 under 35 U.S.C. §103(a) as being unpatentable over Mack, David H., U.S. Patent No. 6,303,301, (hereinafter "Mack"), in view of Mangiameli *et*

al., European J. Operational Res., 93:402-417(1996) (hereinafter "Mangiameli") and Kohonen (Reference AR of Form 1449 filed on July 14, 2000; hereinafter "Kohonen"). The Examiner states that Mack discloses methods of cluster analysis for gene expression monitoring. In particular, the Examiner states that Mack's methods comprise receiving gene expression values of datapoints, clustering the datapoints, and providing output display indicating the cluster of the datapoints. The Examiner further states that while Mack does not explicitly disclose clustering using SOMs, it does suggest using alternative statistical methods. The Examiner states that Mangiameli applied SOM and seven hierarchical methods to 252 messy data sets and found that SOMs are significantly superior in both robustness and accuracy to other clustering methods. Additionally, the Examiner states that Kohonen teaches every aspect of SOMs. Accordingly, the Examiner asserts that one of ordinary skill in the art would have been motivated to modify the method of Mack to use SOMs as suggested and taught by Mangiameli and Kohonen for the cluster analysis of gene expression data to achieve its superiority in accuracy and robustness.

Applicants respectfully disagree. The primary reference, Mack, only generally refers to the use of clustering methods. Mack applies a cluster analysis in a unilateral, class inclusion/exclusion (of a certain class) approach. Mack, however, provides no algorithm or specific description of how the clustering analysis is performed. According to Mack, it uses cluster analysis for determining whether mutations in up-stream regulatory genes exist by monitoring down-stream gene expression. In particular, Mack provides a method for determining whether a down-stream gene expression indicates a p53 mutation or not. The methods in Mack determine whether or not a gene expression falls into a known, pre-determined category. Contrary to this binary (inclusion/exclusion) classification for known classes, the claimed invention advantageously, in an unsupervised learning fashion, classifies gene expression data into *multiple* classes. Importantly, the claimed invention *does not require any information of known classes*. The claimed invention can classify gene expression data into unknown classes, redefine classes, or rediscover classes.

Hence, combining Mack with Mangiameli would be improper because one of skill in the art would not use SOMs, an accurate and robust neural network for messy empirical data as described in Mangiameli, for data requiring only a simple, parametric analysis, like that found in Mack.

Assuming *arguendo* that Mack and Mangiameli were combined, the present invention would not result. Instead one would be motivated to use Mack in a subsequent stage after processing messy empirical data. For example, once the claimed invention is applied to gene expression data to determine what classes exists, including previously unknown classes, then the methods described in Mack could be used to assign a sample to one of those classifications.

Thus no combination of the prior or cited art makes obvious the present invention as now claimed. In order to highlight the foregoing distinctions over the prior art, base claim 1 recites:

> using a self organizing map, clustering the datapoints such that the datapoints that exhibit similar patterns are clustered together into respective clusters in a manner free of predetermined association of patterns with respective clusters.

Claim 11, the other independent claim, recites similar language. Accordingly, Applicants believe that the rejection of Claims 1-18 under 35 U.S.C. §103 is overcome. Reconsideration of Claims 1-18 is respectfully requested.

CONCLUSION

In view of the above amendments and remarks, it is believed that all claims are in condition for allowance, and it is respectfully requested that the application be passed to issue. If the Examiner feels that a telephone conference would expedite prosecution of this case, the Examiner is invited to call the undersigned at (978) 341-0036.

Respectfully submitted,

HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

By Antoinette G. Giugliano

Antoinette G. Giugliano
Registration No. 42,582
Telephone: (978) 341-0036
Facsimile: (978) 341-0136

Concord, MA 01742-9133
Dated: 9/19/02

MARKED UP VERSION OF AMENDMENTS

Specification Amendments Under 37 C.F.R. § 1.121(b)(1)(iii)

Replace the paragraph at page 6, lines 8 through 10 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

Figures 6A-6B summarize the experiments performed under various conditions for a Yeast Cell Cycle analysis. [This summary and all data obtained for the experiments can be found at http://genome-www.stanford.edu/cellcycle.]

Replace the paragraph at page 21, lines 10 through 12 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

Yeast Experiments: Yeast data was downloaded from [http://genome-www.stanford.edu/cellcycle] a Stanford University website. The 90 minute time point was excluded because of difficulties with scaling. See Figures 6A-B.

Replace the paragraph at page 21, line 13 through page 22, line 6 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

Expression Analysis: A detailed protocol [is at http://www.genome.wi.mit.edu/MPR, and pertinent portions of it can also] can be found in Example 5. Briefly, 1 μg mRNA was used to generate first strand cDNA using a T7-linked oligo-dT primer. Following second strand synthesis, *in vitro* transcription (Ambion) was performed with biotinylated UTP and CTP (Enzo), resulting in 40-80 fold linear amplification of RNA. 40 μg of biotinylated RNA was fragmented to 50-150 nucleotide size prior to overnight hybridization to Affymetrix HU6000 arrays. Arrays contain probe sets for 6416 human genes (5223 known genes and 1193 ESTs). Because probe sets for some genes are present more than once on the array, the total number on the array is 7227. Following washing,

arrays were stained with streptavidin-phycoerythrin (Molecular Probes) and scanned on a Hewlett-Packard scanner. Intensity values were scaled such that overall intensity for each chip of the same type was equivalent. Intensity for each feature of the array was captured using GeneChip software (Affymetrix, Inc.), and a single raw expression level for each gene was derived from the 20 probe pairs representing each gene using a trimmed mean algorithm. A threshold of 20 units was assigned to any gene with a calculated expression level below 20, since discrimination of expression below this level could not be performed with confidence.

Replace the paragraph at page 24, lines 12 through 21 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

The myeloid leukemia cell line HL-60, which undergoes macrophage differentiation upon treatment with the phorbol ester TPA was studied. Nearly 100% of HL-60 cells become adherent and exit the cell cycle within 24 hours of TPA treatment. To monitor this process at the transcriptional level, anti-sense cRNA was prepared from cells harvested at 0, 0.5, 4 and 24 hrs after TPA stimulation (see Example 1). Samples were then hybridized to expression-monitoring arrays from Affymetrix, Inc., containing oligonucleotide probes for 5223 known human genes and 1193 expressed sequence tags (ESTs), and hybridization intensities were determined for each gene. [The list of genes on the arrays and all expression data are available at http://www.genome.wi.mit.edu/MPR.]

Replace the paragraph at page 27, lines 13 through 22 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

The present invention was applied to more complex datasets involving multiple cell lines: HL-60 and the similar myeloid cell line U937, which also undergoes macrophage differentiation in response to TPA; Jurkat, a T-cell line that acquires many hallmarks of T-cell activation in response to TPA; and NB4, an acute promyelocytic leukemia cell line that undergoes neutrophilic differentiation in response to all-trans retinoic acid (ATRA). A total of 17 RNA samples were generated, yielding 6416 datapoints in 17-dimensional space. Of these, 1036 genes passed the

variation filter. The genes were classified with a 6x4 SOM (Figure 5A-X), thereby grouping the 1036 genes into 24 categories. [See http://www.genome.wi.mit.edu/MPR for the entire database.]

Claim Amendments Under 37 C.F.R. § 1.121(c)(1)(ii)

1. (Amended) In a computer system, a method for clustering a plurality of datapoints, wherein each datapoint is a series of gene expression values, wherein the method comprises:

a) receiving the gene expression values of the datapoints;

b) using a self organizing map, clustering the datapoints such that the datapoints that exhibit similar patterns are clustered together into respective clusters in a manner free of predetermined association of patterns with respective clusters; and

c) providing an output indicating the clusters of the datapoints.

11. (Amended) In a computer system, a method for grouping a plurality of datapoints, wherein each datapoint is a series of gene expression values, wherein the method comprises:

a) receiving gene expression values of the datapoints;

b) filtering out any datapoints that exhibit an insignificant change in the gene expression value, such that working datapoints remain;

c) normalizing the gene expression value of the working datapoints;

d) using a self organizing map, grouping the working datapoints such that the datapoints that exhibit similar patterns are grouped together into respective clusters in a manner free of predetermined association of patterns with respective clusters; and

e) providing an output indicating the groups of the datapoints.